The International Journal of Biochemistry & Cell Biology 58 (2015) 71–80

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Structure and function of *Mycobacterium smegmatis* 7-keto-8-aminopelargonic acid (KAPA) synthase☆

Shanghua Fan[a,1], De-Feng Li[b,1], Da-Cheng Wang[b], Joy Fleming[b], Hongtai Zhang[b], Ying Zhou[b], Lin Zhou[c], Jie Zhou[d], Tao Chen[c], Guanjun Chen[a,*], Xian-En Zhang[b,**], Lijun Bi[b,**]

[a] *State Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Jinan, Shandong 250100, China*
[b] *National Laboratory of Biomacromolecules and CAS Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China*
[c] *Center for Tuberculosis Control of Guangdong Province, Guangdong 510630, China*
[d] *The 4th People's Hospital, Foshan, Guangdong 528000, China*

ARTICLE INFO

Article history:
Received 15 September 2014
Received in revised form 24 October 2014
Accepted 10 November 2014
Available online 18 November 2014

Keywords:
Mycobacterium smegmatis
Biotin synthesis pathway
7-Keto-8-aminopelargonic acid (KAPA) synthase
Crystal structure
Active site

ABSTRACT

The biotin biosynthesis pathway is an attractive target for development of novel drugs against mycobacterial pathogens, however there are as yet no suitable inhibitors that target this pathway in mycobacteria. 7-Keto-8-aminopelargonic acid synthase (KAPA synthase, BioF) is the enzyme which catalyzes the first committed step of the biotin synthesis pathway, but both its structure and function in mycobacteria remain unresolved. Here we present the crystal structure of *Mycobacterium smegmatis* BioF (MsBioF). The structure reveals an incomplete dimer, and the active site organization is similar to, but distinct from *Escherichia coli* 8-amino-7-oxononanoate synthase (EcAONS), the *E. coli* homologue of BioF. To investigate the influence of structural characteristics on the function of MsBioF, we deleted *bioF* in *M. smegmatis* and confirmed that BioF is required for growth in the absence of exogenous biotin. Based on structural and mutagenesis studies, we confirmed that pyridoxal 5′-phosphate (PLP) binding site residues His129, Lys235 and His200 are essential for MsBioF activity *in vivo* and residue Glu171 plays an important, but not essential role in MsBioF activity. The N-terminus (residues 1–37) is also essential for MsBioF activity *in vivo*. The structure and function of MsBioF reported here provides further insights for developing new anti-tuberculosis inhibitors aimed at the biotin synthesis pathway.

1. Introduction

Biotin is an essential cofactor for enzymes that function in the carboxylation, decarboxylation, and transcarboxylation reactions found in processes such as fatty acid biosynthesis, gluconeogenesis, and amino acid metabolism (Moss and Lane, 1971). While plants and microorganisms can synthesize biotin (Cronan and Lin, 2011), mammals must obtain it from exogenous sources (Said, 2009). Accumulating evidence shows that the *de novo* biotin biosynthesis pathway is essential for the survival and growth of mycobacteria in a variety of living environments (Keer et al., 2000; Sassetti and Rubin, 2003; Rengarajan et al., 2005; Dey et al., 2010; Woong Park et al., 2011; Yu et al., 2011). The biotin biosynthesis pathway has therefore been considered as a potential target for the development of new antibiotics.

The biotin synthesis pathway can divided into two parts, namely biotin pimelate moiety synthesis and four subsequent conserved steps carried out by four committed enzymes (Salaemae et al., 2011; Lin et al., 2010), namely 7-keto-8-aminopelargonic acid synthase (KAPAS or BioF), 7,8-diaminopelargonic acid synthase (DAPAS or BioA), dethiobiotin synthetase (DTBS or BioD) and biotin synthase (BS or BioB). KAPA synthase, the enzyme catalyzing the first committed step of the biotin synthesis pathway, is a pyridoxal 5′-phosphate (PLP)-dependent enzyme belonging to subclass II of the aminotransferase family (Alexeev et al., 1998; Bhor et al., 2006). It has been studied in detail in microbes like *Escherichia coli*

Abbreviations: KAPA, 7-keto-8-aminopelargonic acid; PLP, pyridoxal 5′-phosphate; ADS, albumin-dextrose-saline; TEV, tobacco etch virus; KAM, N-(7-keto-8-aminopelargonic acid)-(3-hydroxy-2-methyl-5-phosphonooxymethyl-pyridin-4-yl-methane); PDB, Protein Data Bank.

☆ *PDB accession number*: The structural factors and coordinates of MsBioF have been deposited in the Protein Data Bank with the PDB ID code 3WY7.

* Corresponding author. Tel.: +86 631 5688199.
** Corresponding authors. +86 10 64888148/464.
E-mail addresses: guanjun@sdu.edu.cn (G. Chen), zhangxe@sun5.ibp.ac.cn (X.-E. Zhang), blj@ibp.ac.cn (L. Bi).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.biocel.2014.11.006
1357-2725/

(Alexeev et al., 1998; Webster et al., 2000; Alexeev et al., 2006), *Bacillus subtilis* (Snapper et al., 1990) and *Bacillus sphaericus* (Bower et al., 1996; Ploux and Marquet, 1996; Ploux et al., 1999). Similarities in the sequences of the enzymes in these organisms and the reactions they catalyze suggest that they share a common catalytic mechanism (Bhor et al., 2006). The crystal structures of KAPA synthase from *E. coli* (PDB code 1BS0, Alexeev et al., 1998) and *Francisella tularensis SCHU S4* (PDB code 4IW7) have been solved. The structure of *E. coli* 8-amino-7-oxononanoate synthase (EcAONS), the *E. coli* KAPA synthase homologue, consists of two protomers, each protomer being composed of three domains, the small N-terminal, large central and C-terminal domains. Its cofactor, pyridoxal 5′-phosphate (PLP), is covalently attached to Lys236 and can be seen in the X-ray structure at the junction between the β-sheets at the bottom of the deep cleft formed between the central and C-terminal domains (PDB code 1DJE, Webster et al., 2000). The other catalytic site residues: His133, Glu175, Asp204 and His207, are located in the central domain, flanking the cleft. Though both the structure and the mechanism of the *E. coli* homologue of BioF, EcAONS, have been well studied, the structure and function of BioF in mycobacteria remains unresolved.

The genome sequence of *Mycobacterium tuberculosis* contains two *bioF* genes (*bioF1*, *Rv1569* or *MtbioF1* and *MtbioF2*, *Rv0032* or *MtbioF2*) (Salaemae et al., 2011). Targeted gene knockout studies of *MtbioF1* in *M. tuberculosis* have provided indirect evidence that there is no redundancy between the two genes. Deletion of *MtbioF1* using a transposon insertion method highly attenuated growth *in vitro* and in mice (Sassetti and Rubin, 2003; Dey et al., 2010). The *Mycobacterium smegmatis* genome encodes only one *bioF* gene (*MSMEG_3189*, *MsbioF*). MsBioF shares 76% identity with MtBioF1 and 35% identity with the BioF domain of MtBioF2. The study of BioA and BioD, the enzymes catalyzing the second and third of the four conserved steps in *M. tuberculosis* biotin synthesis suggests that there are significant differences between the enzymes in the biotin synthesis pathway of mycobacteria and those of other organisms (Dey et al., 2010). In contrast to *B. sphaericus* (Ploux et al., 1999), D-alanine is a substrate rather than a competitive inhibitor of BioF in *M. tuberculosis*, demonstrating the broad substrate specificity of the mycobacterial enzyme (Bhor et al., 2006). We therefore expect that mycobacterial BioF may also be different from that in other organisms.

Here we present the structure of MsBioF determined using the molecular replacement method. To study its structural characteristics, we knocked out the *bioF* gene in *M. smegmatis*. We found that the *bioF* gene is essential for the growth of *M. smegmatis* in the absence of exogenous biotin. Phenotypes of the *M. smegmatis bioF* deletion mutant were complemented by the expression of *M. tuberculosis bioF1* but not *bioF2*, suggesting that BioF plays a conserved role in mycobacterial biotin biogenesis. Based on structural analysis and site-directed mutagenesis, we determined the putative catalytic site residues of MsBioF and confirmed which residues are essential for *M. smegmatis* to grow in the absence of exogenous biotin. Our results also confirmed that dimerization of the two MsBioF protomers via domain swapping is essential for biotin synthesis. Design of inhibitors that prevent domain swapping in the MsBioF dimer may be a suitable strategy for inactivating MsBioF, and therefore inhibiting the biotin synthesis pathway in mycobacterial pathogens.

2. Materials and methods

2.1. Strains, media and culture conditions

M. smegmatis MC2 155 and its derivatives were grown in Sauton's liquid medium containing 0.02% tyloxapol (for growth curve determination), or in Middlebrook 7H9 liquid medium (Difco) supplemented with 10% albumin-dextrose-saline (ADS) (for vector construction), or on Middlebrook 7H11 agar plates containing 10% OADC supplement (Becton Dickinson) and 0.5% glycerol unless otherwise stated. As required, hygromycin B (50 or 150 mg/mL), kanamycin (30 mg/mL), X-Gal (50 μg/mL) and sucrose (2%, w/v) were added to the cultures. Biotin (Sigma), at a concentration of 1 μM, was added as indicated.

2.2. Cloning, expression, purification, and crystallization of MsbioF

The procedures used have been described in detail elsewhere (Fan et al., 2014). Briefly, the coding gene of *MsbioF* was cloned into expression plasmid pET28a (Invitrogen) with His6 and a recombinant TEV protease (tobacco etch virus protease) cleavage site at the N-terminus, and the resulting plasmid, pET28a:*MsbioF*, was then transformed into *E. coli* BL21. MsBioF protein expression was induced with IPTG and MsBioF protein was then purified through sequential chromatographic steps on the following columns: a Ni–NTA (Novagen) column, HiPrep 26/10 Desalting Column, Q-Sepharose anion exchange column and a Superdex 75 10/300 GL column. Initial hanging-drop vapour-diffusion crystallization conditions for the purified protein were screened using Crystal Screen 1 and Crystal Screen 2 (Hampton Research, USA) at 298 K. Crystals suitable for diffraction experiments were obtained within 14 days using the condition 0.05 M DL-Malic acid pH 7.5 containing 20% (w/v) polyethylene glycol 3350.

2.3. Data collection, structure determination and refinement

The procedures used have been described in detail elsewhere (Fan et al., 2014). Prior to the collection of diffraction data, crystals were quickly transferred to corresponding cryoprotectant solutions containing 5% (v/v) glycerol, then flash-frozen in liquid nitrogen. Diffraction data was collected from a single crystal on beamline 17 U of the Shanghai Synchrotron Radiation Facility using an ADSC Q315 CCD detector.

The crystal structure of MsBioF was solved by molecular replacement using the structure of EcAONS (PDB code 1BS0) as the search model. Further model refinement was carried out using PHENIX. Data collection and refinement statistics are given in Table 1.

2.4. MsbioF deletion

The *M. smegmatis MsbioF* deletion strain (*ΔMsbioF strain*) was constructed using the unmarked deletion strategy described by Parish and Stoker (2000), except that we used overlap PCR rather than restriction endonuclease digestion to construct a mutant allele. Briefly, to generate the mutant allele, the 3′-portion of *MsbioF* and its upstream region, and the 5′-portion of *MsbioF* and its downstream region were amplified by PCR. We then fused these two ~1200 bp amplicons by overlap extension PCR to generate the mutant allele. The mutant allele was digested with *Not I* and *Xho I* and cloned into a similarly digested p1NIL plasmid (a kind gift from Tanya Parish) resulting in plasmid p1NIL-MA. The *Pac I* cassette from pGOAL19 (hyg, lacZ, P_{hsp60}-sacB) was then cloned into the single *Pac I* site of p1NIL-MA to generate the suicide delivery vector p1NIL-MM which contained the complete *MsbioF* knockout cassette. To enhance the transformation of electrocompetent *M. smegmatis* cells, plasmid DNA p1NIL-MM was treated with NaOH/EDTA (0.2 mM/0.2 mM) (Hinds et al., 1999). Transformants were selected on hygromycin/kanamycin/X-Gal plates and single cross-over transformants were blue. Single cross-over transformants were streaked out into 7H9 liquid culture, resuspended and incubated for 3–4 h. We then plated part of the incubated

Table 1
Data collection, structure determination, and refinement statistics for MsBioF.

Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	70.51, 91.56, 108.71
α, β, γ (°)	90.0, 97.78, 90.0
Wavelength (Å)	1.5418
Resolution (Å)	40.12–2.30 (2.42–2.30)
R_{merge} † (%)	11.9 (37.6)
I/σ	7.4 (2.7)
Refinement statistics	
Resolution (Å)	40.12–2.30
No. reflections	60,733
R_{work}/R_{free}	0.242/0.277
Protein atoms	10,302
Water molecules	411
B-factors (Å^2)	
Protein	40.6
Water	36.9
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.06
Ramachandran plot	
Core region (%)	97.31
Allowed region (%)	2.11
Generous region (%)	0
Disallowed region (%)	0.58
Average *B*-factor (Å)	40.5

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of reflection *hkl*. $\sum_i$ is the sum over the individual measurements of a reflection and $\sum_{hkl}$ is the sum over all reflections.

culture onto sucrose X-Gal plates. The white colonies, all double cross-overs, were Kan^S and resistant to sucrose. PCR and Western blotting were used to distinguish wild-type and mutant cells.

All gene disruptions were tested by PCR using flanking DNA sequences as primers and were further confirmed by DNA sequencing of the PCR products.

2.5. Complementation plasmid and merodiploid strain construction

To complement the *ΔMsbioF* strain, the coding sequence of *MsbioF* was PCR-amplified from *M. smegmatis* MC2 155 genomic DNA using primers *MsbioF*-F and *MsbioF*-R, which contain *BamH I* and *Hind III* restriction sites, respectively. It was then digested by *BamH I* and *Hind III* and cloned into a similarly digested pMV261 plasmid. For heterologous complementation with *M. tuberculosis MtbioF1*, the gene was amplified from *M. tuberculosis* H37Rv genomic DNA using primers *MtbioF1*-A and *MtbioF1*-B and cloned into pMV261. The *MtbioF2* gene was also amplified and cloned into vector pMV261. To identify the active site of MsBioF, site-directed mutagenesis was performed at putative catalytic site residues His129, Cys136, Glu171, Asp200, His203 and Lys235 of MsBioF. All mutations were performed according to the method of Shenoy and Visweswariah (2003). To investigate the role of the α1 helix of the N-terminal domain *in vivo*, we constructed two expression vectors containing MsBioF N-terminal truncation mutants with deletions of residues 1–37 and 1–22, respectively. All plasmids constructed are summarized in Table S1.

The resulting vector was transformed into the *M. smegmatis ΔMsbioF* strain by electroporation, and transformants were selected using kanamycin. Protein levels in cell extracts from all strains were analyzed by SDS-PAGE (100 mg of cell extracts were loaded on the gel) followed by Western blotting performed using polyclonal anti-MsBioF according to standard procedures (Towbin et al., 1979). All strains constructed are summarized in Table S1.

2.6. Antibody production and Western blotting

MsBioF was first purified using the above method, except that the N-terminal 6xHis tag was removed with TEV Protease before nickel column purification. MsBioF antiserum was then obtained by immunizing mice with purified recombinant MsBioF. To detect MsBioF, its homologue MtBioF, and their mutants, 1 mL of *M. smegmatis* stationary phase cells was harvested. Cell pellets were resuspended in 50 mL Laemmli sample buffer and then boiled at 363 K for 30 min. Samples were subjected to 12% SDS-PAGE, followed by transfer to a nitrocellulose membrane. The nitrocellulose membrane was blocked in 1 × PBST with 5% (w/v) nonfat milk for 1 h at room temperature and then incubated overnight at 4 °C with a 1:500 dilution of MsBioF antiserum. After washing the membrane three times with PBST, bound primary antibody was detected by incubating with a secondary HRP-linked anti mouse IgG (1:5000, GE, Healthcare) and chemiluminescent substrate (ECL-plus substrate, GE, Healthcare) according to the manufacturer's instructions.

2.7. Determination of the activity of mycobacterial BioF and its mutants in vivo

To identify whether the expression of BioF and its mutants is required for the growth of *M. smegmatis* MC2 155 in liquid medium, the wild-type strain of *M. smegmatis* MC2 155, the mycobacteral *bioF* mutant and merodiploid strains were grown in 5 mL 7H9 culture containing ADC for 2–3 d at 37 °C until stationary phase. Then a 1 mL culture was centrifuged (6000 rpm, 5 min, 4 °C) and resuspended in 1 mL Sauton's medium supplemented with tyloxapol, and this procedure was repeated three times. 30 μL of the final suspension was used to inoculate flasks containing 15 mL of Sauton's medium supplemented with tyloxapol, with or without 1 μM biotin. Cultures were incubated at 37 °C and the absorption of 500 μL culture samples was determined at λ600 nm every 3 or 6 h.

3. Results

3.1. Crystal structure of MsBioF

The three-dimensional structure of apo MsBioF was determined to 2.3 Å resolution with an $R_{crystal}$ of 0.242 and an R_{free} of 0.277. Refinement statistics are summarized in Table 1. Each asymmetric unit of the cell contains four monomers. The structure indicates that the enzyme forms an incomplete symmetric homodimer (protomers A and B) within the crystal (Fig. 1A), consistent with the structure of homologous protein EcAONS (Alexeev et al., 1998). The 6X-histidine tag and residues 1–37 could not be traced in the electron density map of protomer B of the 404 residue full-length MsBioF, presumably due to flexibility in residues 1–37. In protomer A, however, only the histidine tags and the N-terminal regions (residues 1–10) could not be seen. The motif of residues 1–37 corresponds to the small N-terminal domain of EcAONS (Fig. 1C). Though we could not trace the His-tag and the N-terminus (residues 1–10 in protomer A and residues 1–37 in protomer B) we observed enough space in the crystal to accommodate them and propose that the N-terminal His-tag does not interfere with crystal packing, or the formation of the intermediate dimer described in the following section (Carson et al., 2007).

The structure of protomer B solved here consists of two domains centered on β sheets, corresponding to the larger central domain and the C-terminal domain of EcAONS, respectively (Fig. 1B). By contrast, protomer A not only includes the larger central domain and C-terminal domain but also an N-terminal helix (residues 11–23) which is separated from the central domain by a loop

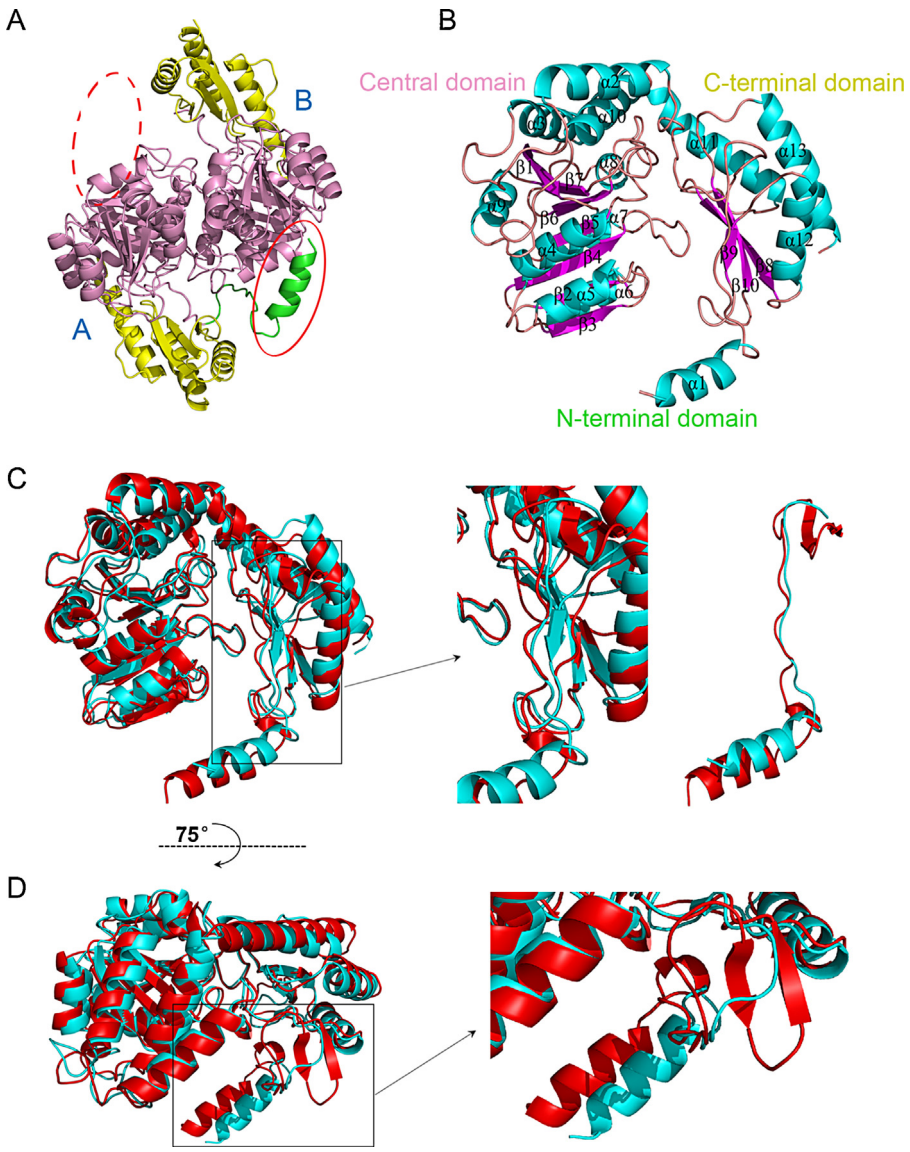

Fig. 1. Overall structure of MsBioF and superposition of MsBioF on EcAONS. (A) Overall structure of the MsBioF dimer consists of protomers A and B. The green helix in the red oval belongs to protomer A and forms the intermolecular strap between protomer A and B. The red dashed oval represents the location where the N-terminal helix belonging to protomer B should be. (B) Protomer structure and secondary structure of MsBioF. Helices, strands and loops are in cyan, magenta and salmon, respectively; (C) Overall and N-terminal domain structure superposition of MsBioF on EcAONS with two views 75° apart around the horizontal. The structure of MsBioF and EcAONS are in cyan and red, respectively. (For interpretation of the references to color in figure legend, the reader is referred to the web version of the article.)

(residues 24–37) and forms an intermolecular strap connecting the larger central domain of protomer A to protomer B (Fig. 1A). The larger central domain and C-terminal domain are composed of amino acid residues 38–283 and 284–382, respectively. The larger central domain is the core and major domain, and is formed from six parallel β strands, β1, β6, β5, β4, β2 and β3, with the antiparallel strand β7 interposed between β1 and β6. The sheet is surrounded by alpha helices α3–α10, with helix α2 being on the periphery (Fig. 1B). Helices α2, α4, α5, α9 and α10 of protomers A and B form the contact surface of the dimer (Fig. 1A). The C-terminal domain consists of a three-stranded antiparallel β sheet formed by β strands β8, β9, and β10, and three helices (α11–α13) which lie against the β-strands (Fig. 1B). The central domain and C-terminal domain form a deep cleft. As previously reported (Webster et al., 2000; Alexeev et al., 1998), the cofactor PLP binds to Lys235 at the bottom of the cleft.

Although there is only 37% sequence identity between MsBioF and EcAONS, their three dimensional structures are almost the same. Superposition of the MsBioF and *E. coli* AONS structures (PDB code 1BS0) reveals one significant difference in their overall structure (Fig. 1C). The residues connecting the α1 of the N-terminal small domain and the central domain form a two-stranded β-sheet in EcAONS. However, the corresponding segment in MsBioF is shorter than that in EcAONS and cannot fold into a two-stranded β-sheet, but rather forms a loop (Fig. 1C). This loop is flexible and may contribute to the disorder of N-terminal helix α1 in protomer B. In addition, the flexible loop could disturb the dimerization of MsBioF via domain swapping of helix α1.

3.2. *Sequence alignment and structure superposition reveals the putative catalytic site*

The binding of cofactor PLP and substrate binding are key steps in the MsBioF catalytic mechanism. However, we failed to co-crystallize MsBioF and PLP/substrate. Our mass spectrometry

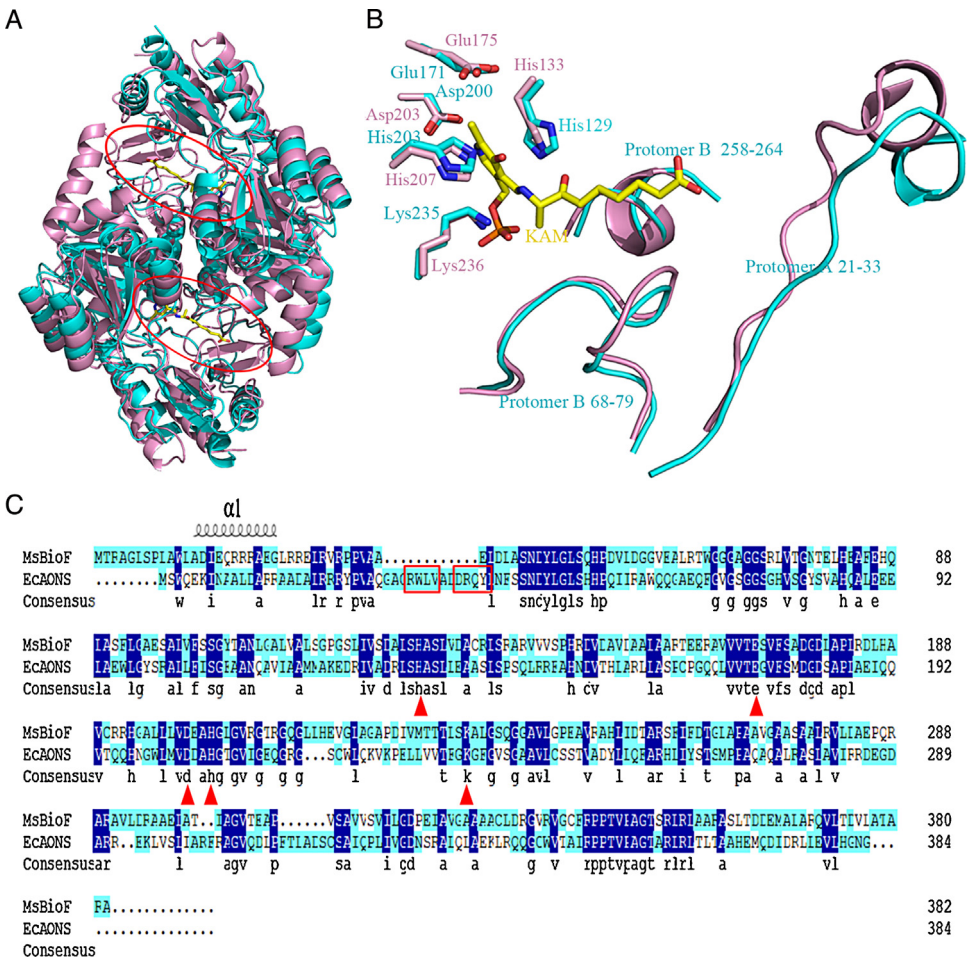

Fig. 2. Sequence alignment and structure superposition reveal the putative catalytic site. (A) Structure superposition of the MsBioF dimer (cyan) on the EcAONS dimer (pink) in complex with the external aldimine intermediate 8-amino-7-oxononanoate-PLP (KAM) (yellow, shown in the red oval). Each promoter of the dimer contributes an active site to bind KAM, which lies at the cleft. (B) Close-up view of the superposition of the active site region of MsBioF protomer A on that of EcAONS in complex with the external aldimine intermediate 8-amino-7-oxononanoate-PLP (KAM) (yellow). Residues (all belonging to protomer A) and secondary elements of MsBioF and EcAONS are in cyan and pink, respectively. (C) Alignment of amino acid sequences of MsBioF and EcAONS. The putative catalytic site residues are indicated by red triangles; Residues in the two black boxes form the two-stranded β-sheet in EcAONS between α1 of the N-terminal small domain and the central domain in Fig. 1; α1 of the N-terminal small domain is also shown. (For interpretation of the references to color in figure legend, the reader is referred to the web version of the article.)

experiments did not demonstrate the formation of a covalent complex between PLP and MsBioF. Binding between MsBioF and cofactor PLP was too weak to be detected in ITC experiments, suggesting that MsBioF does not readily react with PLP to form a covalent complex *in vitro*. We thus used homologue modeling to study the interaction between MsBioF and PLP, as well as its substrates.

The catalytic reaction mechanism of 7-keto-8-aminopelargonic synthase (EcAONS) in *E. coli* has been elaborated in detail (Webster et al., 2000; Alexeev et al., 1998, 2006). 7-Keto-8-aminopelargonic acid-PLP (KAM) is an external aldimine intermediate product of the reaction and the structure of the KAM-AONS complex has also been resolved (PDB code 1DJ9, Webster et al., 2000). Here we modeled 7-keto-8-aminopelargonic acid-PLP (KAM) into the structure of MsBioF by aligning it with the homologous structure 1DJ9 (Webster et al., 2000) containing KAM. As shown in Fig. 2A, both subunits of the 'catalytic dimer' contribute to an active site to bind KAM, which lies at the subunit and domain interface. The PLP head of KAM is located in the bottom of the cleft between the large domain and the C-terminal domain of one protomer. The long tail of pimeloyl-CoA lies in the cavity formed by the small N-terminal domain and the C-terminal domain of one monomer and the large domain of the other. Combining structural information with that obtained from sequence alignment of MsBioF and EcAONS, we identified five conserved residues in MsBioF, His129, Glu171, Asp200, His203 and Lys235, corresponding to residues His133, Glu175, Asp204, His207, and Lys236 in EcAONS, as putative catalytic site residues (Fig. 2B and C). All these residues are located on the loops of the large domain except for Glu171 which lies at the terminus of β4. These residues interact with the PLP head of KAM, and the long carbon chain of KAM is surrounded by three loops, residues 21–33 of one protomer (A), and residues 72–83 and 258–264 of the other protomer (B) (Fig. 2B). According to previous reports, atom NZ of Lys235 in the putative active site can covalently bind atom C4A of cofactor PLP, the carboxylate of Asp200 is hydrogen-bonded to atom N1 of PLP, and atom O3 of PLP is hydrogen-bonded to His203 (Alexeev et al., 1998). Glu171 is negatively charged and pairs with Ser175 (Ser179 in EcAONS) to remove the proton from the O3 of PLP by interacting with His133, and His207 (Webster et al., 2000). The location and orientation of Glu171, Asp200, His203, and Lys235 are very similar to those in EcAONS. However, residue His129 has two different conformations differing from each other mainly by main chain conformation and imidazole ring orientation: the imidazole ring of the His129 of one protomer faces in the opposite direction of the cleft, while the His129 of the other protomer faces in the same direction as that in EcAONS. As His133 in EcAONS forms a parallel, stacked arrangement with the PLP ring via the side chain and plays an important role in the mechanism (Webster et al., 2000; Alexeev

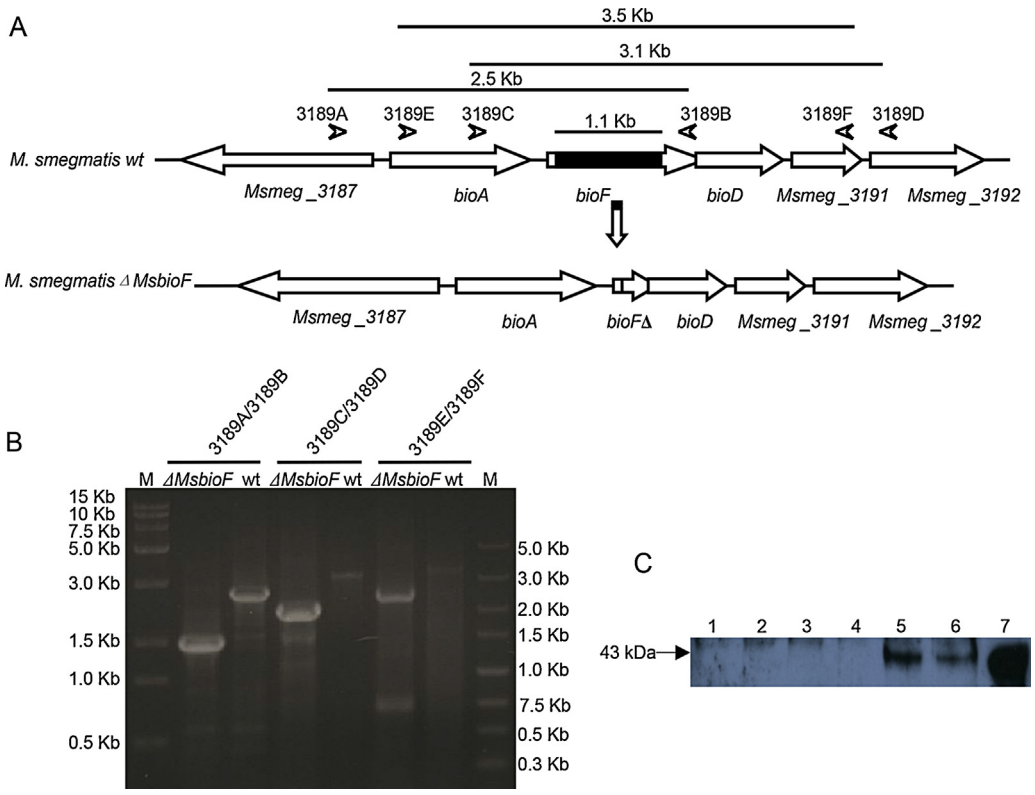

Fig. 3. Construction of *M. smegmatis ΔMsbioF* and complementation plasmids. (A) Schematic diagram of the genomic organization of *MsbioF* and its adjacent genes before and after *MsbioF* deletion. Positions of primers used for verifying the mutant strain are indicated by arrowheads and the expected sizes of PCR products are indicated. (B) PCR analysis of the wild type *M. smegmatis* and *ΔMsbioF* strains using various combinations of specific primers. (C) Western blotting of MsBioF expression in the *M. smegmatis wt* strain and *ΔMsbioF* strains with an anti-MsBioF antiserum. Lanes 1–4, *ΔMsbioF* strains; Lanes 5–6, *wt*; Lane 7, control: purified MsBioF.

et al., 1998), we speculate that the side chain (an imidazolyl group) conformation of His129 in MsBioF may be regulated by its adjacent residues, translocating from the former to the latter conformation, and is then involved in PLP recognition, in a manner similar to that described for EcAONS.

3.3. *MsbioF is required for the growth of M. smegmatis in the absence of exogenous biotin*

The MsBioF coding gene is located between *MsbioA* (*MSMEG_3188*) and *MsbioD* (*MSMEG_3190*), and there is a 4 bp

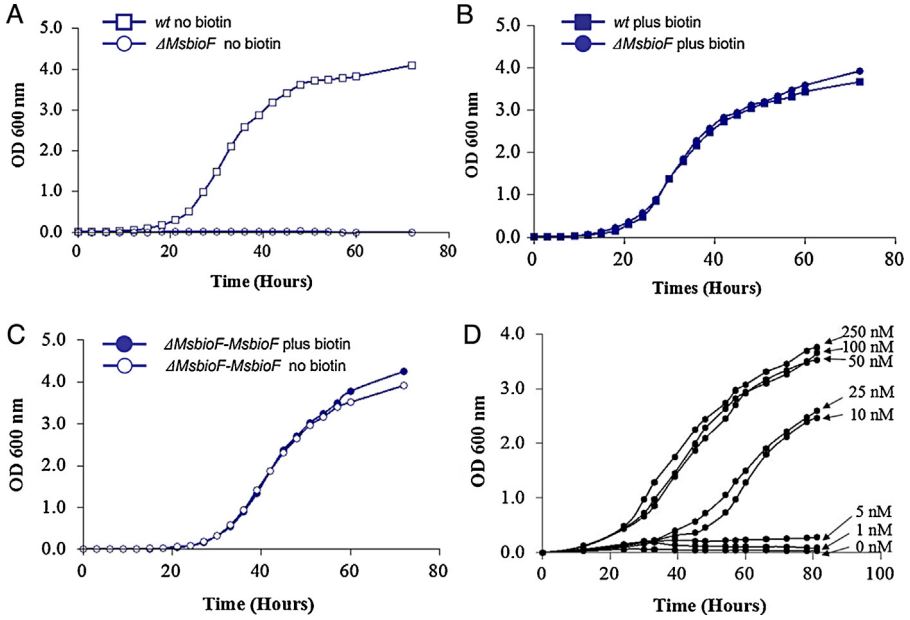

Fig. 4. *MsbioFs* is required for the growth of *M. smegmatis* in the absence of exogenous biotin. (A) Growth of the *wt M. smegmatis* and *ΔMsbioF* strains in the absence of exogenous biotin. (B) Growth of the *wt M. smegmatis* and *ΔMsbioF* strains with exogenous biotin. (C) Growth of complementary strain *ΔMsbioF-MsbioF*. (D) Growth of the *ΔMsbioF* strain in Sauton's medium with varying concentrations of biotin. Open squares and circles represent data in the absence of exogenous biotin, while closed squares and circles represent data in the presence of exogenous biotin. Data presented are from individual cultures and are representative of three independent experiments.

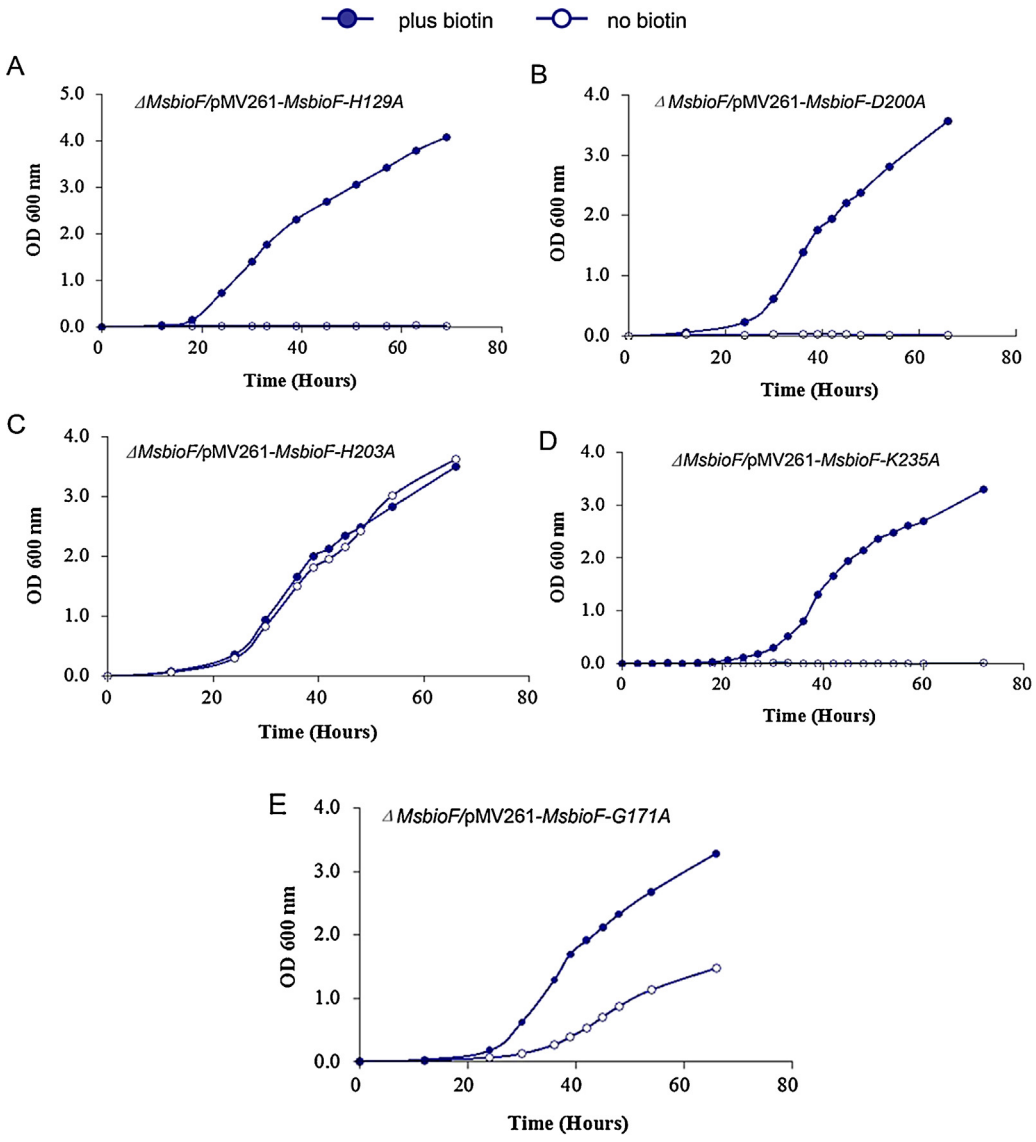

Fig. 5. Key residues of MsBioF involved in biotin synthesis as determined by site-directed mutagenesis. (A) Growth of complementary strain *ΔMsbioF-MsbioF-H129A*. (B) Growth of complementary strain *ΔMsbioF-MsbioF-D200A*. (C) Growth of complementary strain *ΔMsbioF-MsbioF-H203A*. (D) Growth of complementary strain *ΔMsbioF-MsbioF-K235A*. Open circles represent data in the absence of exogenous biotin, and closed circles represent data in the presence of exogenous biotin. Data presented are from individual cultures and are representative of three independent experiments.

overlap between *bioF* and *bioD*. To maintain the integrity of the other genes, we deleted the intermediary 1083 bp section of the *MsbioF* gene, generating a *MsbioF* unmarked mutant (*ΔMsbioF*) using a two-step cross-over method (Fig. 3A) (Parish and Stoker, 2000). The *ΔMsbioF* mutant strains were confirmed by PCR (Fig. 3B), DNA sequencing and Western blotting (Fig. 3C).

We cultured the wild type *M. smegmatis* strain and the *ΔMsbioF* strain in Sauton's medium with and without biotin. As expected, in the presence of 1 μM exogenous biotin, the wild type strain and the *ΔMsbioF* strain both grew with a similar growth curve. In the absence of exogenous biotin, the wild type strain still grew, but the *ΔMsbioF* strain did not (Fig. 4A and B). To further confirm that the inability of the *ΔMsbioF* strain to grow in the absence of exogenous biotin was actually caused by knockout of *MsbioF*, we constructed pMV261: *MsbioF*, a complementary plasmid containing the *MsbioF* gene, using Western blotting to show that it could express MsBioF. Results showed that, like the wild type, the complemented mutant strain could grow in the presence or absence of biotin (Fig. 4C). Together, these results show that *MsbioF* plays an essential role in the growth of *M. smegmatis* in the absence of exogenous biotin. In addition, we determined the minimal concentration of biotin required for *ΔMsbioF* to grow at the same rate as the wild type. When the concentration of biotin was lower than 5 nM, *ΔMsbioF* showed little or no growth (Fig. 4D). The growth of *ΔMsbioF* reached wild type level only when the biotin concentration was higher than 50 nM, which is at least 5-fold greater than the concentration of biotin in normal human plasma (Mock and Malik, 1992). These observations on *M. smegmatis*, together with previous reports on *M. tuberculosis* (Sassetti and Rubin, 2003; Dey et al., 2010), indicate that *de novo* synthesis of biotin is required for both fast-growing and slow-growing mycobacteria.

3.4. Site-directed mutagenesis identified the key residues of MsBioF involved in biotin synthesis

To determine which of the putative catalytic site residues are essential for the activity of MsBioF *in vivo*, we constructed complementary plasmids containing MsBioF site-directed mutants in which these residues were substituted with alanine. We then compared the growth of these merodiploid strains containing site

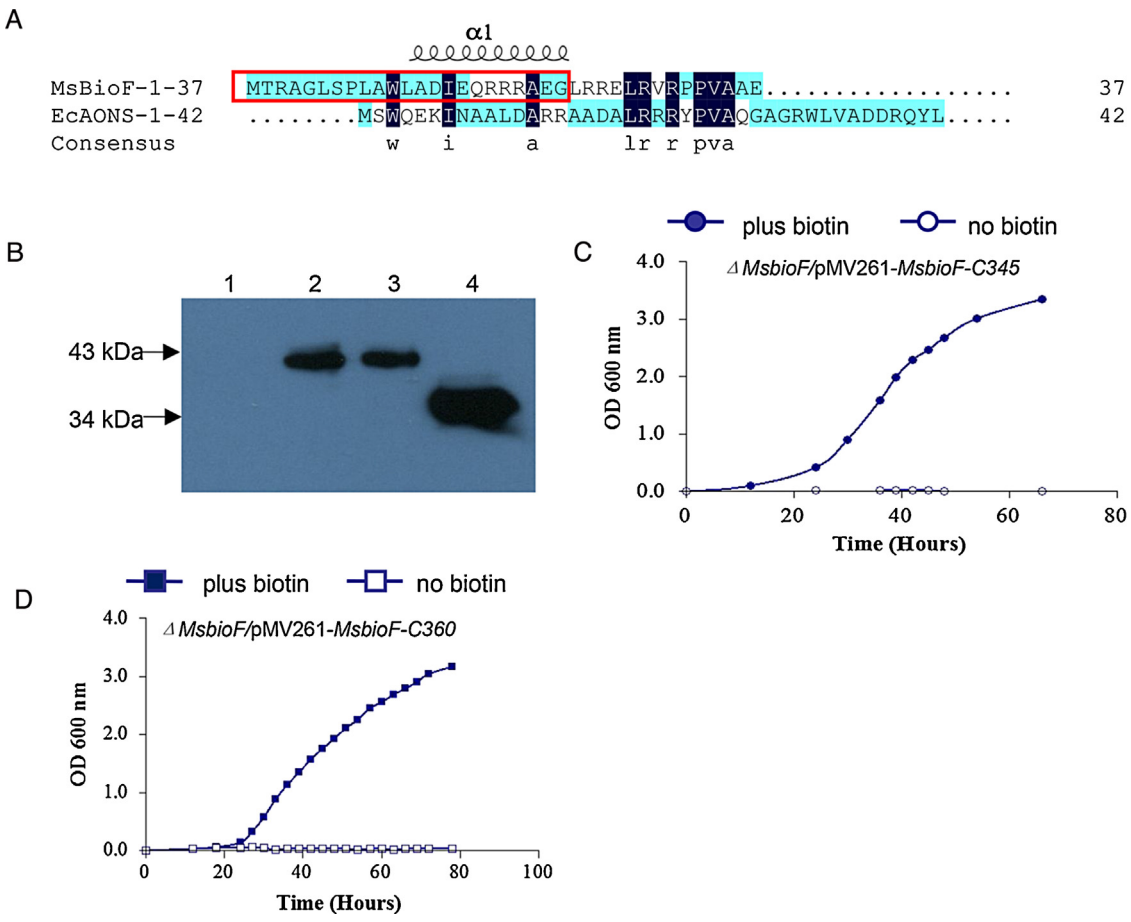

Fig. 6. Dimerization of MsBioF via domain swapping is essential for biotin synthesis. (A) Alignment of the amino acid sequence of the small N-terminal domain of MsBioF with that of AONS. Residues in the red box represent the deleted sequence (residues 1–22) in *ΔMsbioF-MsbioF-C360*. The sequence (1–37) of MsBioF is truncated in *ΔMsbioF-MsbioF-C345*. (B) Western blot of the expression of MsBioF and its truncation mutants lacking residues 1–37. Lane 1, *ΔMsbioF* whole cell lysate; Lane 2, *wt M. smegmatis* whole cell lysate; Lane 3, purified MsBioF; Lane 4, *ΔMsbioF*/pMV261-*MsbioF-C345* whole cell lysate; (C) Growth of complementary strain *ΔMsbioF-MsbioF-C345* which expresses only the 345 residues of the MsBioF C-terminal. (D) Growth of complementary strain *ΔMsbioF-MsbioF-C360* which expresses only the 360 residues of the MsBioF C-terminal. Open circles and squares represent data in the absence of exogenous biotin, and closed circles and squares represent data in the presence of exogenous biotin. Data presented are from individual cultures and are representative of three independent experiments. (For interpretation of the references to color in figure legend, the reader is referred to the web version of the article.)

mutants with that of the wild type *M. smegmatis* strain. Mutation of His129 (Fig. 5A), Asp200 (Fig. 5B) and Lys235 (Fig. 5D) resulted in complete loss of growth in merodiploid strains, suggesting that these residues are essential for the role of MsBioF in the biotin synthesis pathway. Mutation of Glu171 led to growth attenuation in *M. smegmatis* (Fig. 5E), while mutation of His203 (Fig. 5C) had no influence on growth. These results suggest that His129, Asp200 and Lys235 are required for the activity of MsBioF, that Glu171 has a significant influence on, but is not essential for MsBioF activity, and that His203 does not play an important role in biotin synthesis.

3.5. Dimerization of MsBioF via domain swapping is essential for biotin synthesis

The structure of EcAONS shows that its dimer is formed mainly via swapping of the small N-terminal domain (Alexeev et al., 1998). In the structure of MsBioF, only the small N-terminal domain of protomer A interacted with protomer B, the N-terminal domain of which could not be traced. To determine the role of the small N-terminal domain in the activity of MsBioF, we analyzed the activity of MsBioF truncation mutants lacking residues 1–37 or 1–22 (Fig. 6A). *ΔMsbioF* was not complemented by the expression of either of these two MsBioF truncation mutants (Fig. 6C and D), indicating that the small N-terminal domain, i.e. α1, is essential for the activity of MsBioF. These results, together with a previous report that the small N-terminal domain of homologous protein EcAONS plays an important role in dimer formation (Alexeev et al., 1998), led us to conclude that the active state of MsBioF *in vivo* is a dimer. Analyzing the dimer structure of MsBioF, we found that only one small N-terminal domain can be seen swapping with the other monomer. This may be because the structure resolved here is an intermediate state in which domain swapping cannot be completely visualized due to the flexibility of the loop between α1 and the central domain. In addition, the loop of the small N-terminal domain, corresponding to residues 23–37 in MsBioF, may be involved in binding the substrate in the cleft (Alexeev et al., 1998) after dimerization of EcAONS. Therefore, residues 1–37 of MsBioF contribute to MsBioF dimerization via residues 1–22 and to substrate binding via residues 23–37 and are important for MsBioF activity in biotin synthesis.

3.6. The role of MsbioF in M. smegmatis growth can be replaced by MtbioF1 but not by MtbioF2

Previous studies have shown that *MtbioF1* is essential not only for *in vitro* growth but also for growth and establishment of infection *in vivo* (Minnikin et al., 2002; Sassetti and Rubin, 2003; Woong Park et al., 2011). As in *M. tuberculosis*, the *MsbioF* knock-out strain

(*ΔMsbioF*) did not grow in liquid media in the absence of biotin and could be rescued with exogenous biotin (Fig. 1A and B). In addition, all mycobacterial species contain BioF. *M. smegmatis* contains only one copy of the *bioF* gene (*MSMEG_3189*) while *M. tuberculosis* contains two copies (*bioF1* (*Rv1569*) and *bioF2* (*Rv0032*)) (Salaemae et al., 2011). Though there is a high level of active residue sequence identity between *MtbioF1* and *MtbioF2*, only *MtbioF1* has been shown to play an essential role in *M. tuberculosis* growth *in vitro* and *M. tuberculosis* infection and survival in mice (Dey et al., 2010; Sassetti and Rubin, 2003). This indicates that there is no redundancy between *MtbioF1* and *MtbioF2*. The *M. smegmatis* protein MsBioF protein shares 76% identity with the *M. tuberculosis* protein MtBioF1 and 35% identity with the BioF domain of MtBioF2, and active sites are conserved in the three proteins. Here, to determine if the role of BioF is conserved across the mycobacteria, we assessed whether *MtbioF1* and *MtbioF2* could complement for the deletion of *MsbioF*. Complementation with *MtbioF1* removed the dependence of *M. smegmatis ΔMsbioF* on exogenous biotin (Fig. 7A), but complementation with *MtbioF2* was unable to do so (Fig. 7B). We constructed a MtBioF2 truncation lacking the additional acetyltransferase domain that is present in MtBioF2 but not MtBioF1 and observed its BioF activity (Fig. 7C). Results showed that the MtBioF2 truncation was also not able to complement *ΔMsbioF*, suggesting that MtBioF2 is unlikely to play a role in biotin synthesis. BioF thus plays a conserved role in mycobacterial biotin synthesis.

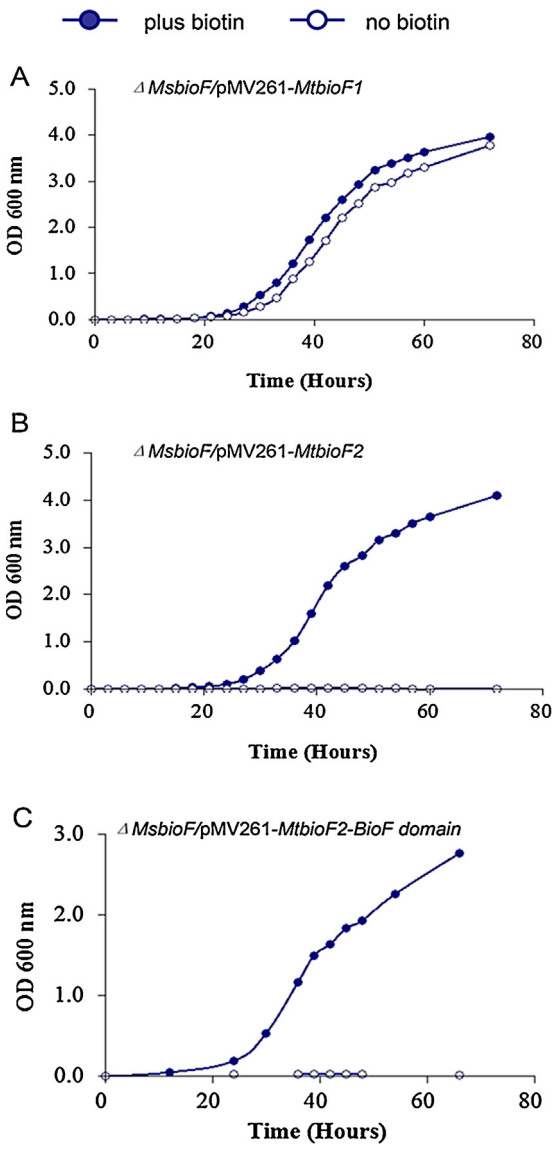

Fig. 7. The role of MsBioF in *M. smegmatis* survival can be replaced by MtBioF1 but not by MtBioF2. (A) Growth of complementary strain *ΔMsbioF-MtbioF1*. (B) Growth of complementary strain *ΔMsbioF-MtbioF2*. (C) Growth of complementary strain *ΔMsbioF-MtbioF2-BioFdomain*. Open circles represent data in the absence of exogenous biotin, and closed circles represent data in the presence of exogenous biotin. Data presented are from individual cultures and are representative of three independent experiments.

4. Discussion

In this study, we have determined the structure of MsBioF, the KAPA synthase enzyme of the *M. smegmatis* biotin synthesis pathway, and investigated its function by generating an *MsbioF* deletion strain (*ΔMsbioF*). A recent report on the structures of BioA and BioD in *M. tuberculosis* suggests that there are significant differences between *M. tuberculosis* and other organisms in enzymes involved in the biotin synthesis pathway (Dey et al., 2010). Our study has confirmed that there are also distinct differences between MsBioF and EcAONS, the homologue of MsBioF in *E. coli*. We have shown that the key MsBioF residues involved in biotin synthesis *in vivo* (Fig. 5), for example, are different from those of EcAONS (Alexeev et al., 1998). Alexeev et al. (1998) speculated that His207 in EcAONS may play an important role in substrate binding and catalysis through formation of a hydrogen-bond with atom O3 of PLP. However, in our study, mutation of His203, the residue corresponding to His207 in MsBioF, had no influence on the growth of *M. smegmatis* (Fig. 5C). In contrast to His133 of EcAONS, His 129, the corresponding residue in MsBioF, has two conformations which may be regulated by its adjacent residues. In addition, residues 23–37, which connect the α1 of the N-terminal small domain and the central domain, form a flexible loop in MsBioF, while the corresponding segment in EcAONS folds into a two-stranded β-sheet (Fig. 1C). The flexibility of this loop made it difficult to resolve the complete structure of protomer B, resulting in the structure of MsBioF presented here being an incomplete dimer. To identify the role of the N-terminal domain, we complemented *ΔMsbioF* with MsBioF truncation mutants lacking residues 1–37 or 1–22 (Fig. 6A). Results confirmed that the N-terminal domain, especially helix α1of MsBioF, is essential for biotin synthesis (Fig. 6C and D). This, together with a previous report that the N-terminal domain plays an important role in the dimerization of EcAONS (Alexeev et al., 1998), led us to conclude that dimerization of MsBioF via domain swapping is required for MsBioF activity. Our functional experiments on MtBioF1 (Fig. 2A), in agreement with sequence alignments (Fig. S1), indicate that the function and structure of BioF is conserved among the mycobacteria. Although amiclenomycin, a natural product isolated from *Streptomyces lavendulae* (Mann and Ploux, 2006), can inhibit the growth of *M. tuberculosis in vitro* through inhibiting BioA (Okami et al., 1974), it has been shown to be inactive in a mouse model of *M. tuberculosis* (Kitahara et al., 1975), possibly due to its poor pharmacokinetic behavior (Woong Park et al., 2011). The structural and functional characteristics of mycobacterial BioF elaborated here thus provide two new insights for the design of inhibitors of the biotin synthesis pathway in mycobacteria. Firstly, inhibitors which target key residues (His129, Asp200 and Lys235) of the active site could be designed to prevent PLP binding. It has been reported that although PLP can covalently attach to Lys236 in the active site of EcAONS, the enzyme is apt to lose the PLP cofactor during crystallization (Webster et al., 2000; Alexeev et al., 1998). Our results indicate that this phenomenon also exists in MsBioF. Based on our study and a previous report on EcAONS, we propose that bonding between all KAPA synthase enzymes and cofactor PLP is weak. Designing inhibitors to prevent the binding of PLP to mycobacterial BioF should thus be

considered. Secondly, anti-mycobacterial drugs could be designed to interfere with the dimerization of BioF. Only the N-terminal domain of one protomer was observed here in the MsBioF dimer structure, suggesting that helix α1 is wobbly and the dimerization of MsBioF through domain swapping is prone to instability. Therefore, anti-mycobacterial drugs designed to target dimerization of mycobacterial BioF would likely inhibit the biotin synthesis pathway. One potential strategy would be to design a hydrocarbon stapled peptide drug (Walensky and Bird, 2014; Walensky et al., 2004) to inhibit the dimerization of MsBioF by competitively binding the central domain and N-terminal domain.

In summary, this structural study is the first to solve the structure of a mycobacterial KAPA synthase and demonstrates that pyridoxal 5′-phosphate (PLP) binding site residues His129, Lys235 and His200 are essential for MsBioF activity *in vivo*, and that dimerization of MsBioF via domain swapping is required for biotin synthesis in *M. smegmatis*. We further validated that *de novo* biotin synthesis is required in *M. smegmatis*. Our study of the structure and function of MsBioF provides new insights into the role mycobacterial BioF plays in biotin synthesis and also open up avenues for the development of novel inhibitors against mycobacterial pathogens such as *M. tuberculosis*.

Acknowledgments

This work was supported by the Chinese Ministry of Health (Grant No: 2013ZX10003006), the Chinese Ministry of Science and Technology (973 Program Grant Nos: 2011CB910300, 2013CB911501), the Chinese Academy of Sciences (Grant No: KSZD-EW-Z-006) and the National Natural Science Foundation of China (Grant No: 31400127).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biocel.2014.11.006.

References

Alexeev D, Alexeeva M, Baxter RL, Campopiano DJ, Webster SP, Sawyer L. The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. J Mol Biol 1998;2:401–19.

Alexeev D, Baxter RL, Campopiano DJ, Kerbarh O, Sawyer L, Tomczyk N, et al. Suicide inhibition of alpha-oxamine synthases: structures of the covalent adducts of 8-amino-7-oxononanoate synthase with trifluoroalanine. Org Biomol Chem 2006;7:1209–12.

Bhor VM, Dev S, Vasanthakumar GR, Kumar P, Sinha S, Surolia A. Broad substrate stereospecificity of the *Mycobacterium tuberculosis* 7-keto-8-aminopelargonic acid synthase: spectroscopic and kinetic studies. J Biol Chem 2006;35:25076–88.

Bower S, Perkins JB, Yocum RR, Howitt CL, Rahaim P, Pero J. Cloning, sequencing, and characterization of the *Bacillus subtilis* biotin biosynthetic operon. J Bacteriol 1996;178:4122–30.

Carson M, Johnson DH, McDonald H, Brouillette C, Delucas LJ. His-tag impact on structure. Acta Crystallogr Sect D: Biol Crystallogr 2007;63:295–301.

Cronan JE, Lin S. Synthesis of the α,ω-dicarboxylicacid precursor of biotin by the canonical fatty acid biosynthetic pathway. Curr Opin Chem Biol 2011;15:407–13.

Dey S, Lane JM, Lee RE, Rubin EJ, Sacchettini JC. Structural characterization of the *Mycobacterium tuberculosis* biotin biosynthesis enzymes 7,8-diaminopelargonicacid synthase and dethiobiotin synthetase. Biochemistry 2010;49:6746–60.

Fan S, Li D, Fleming J, Hong Y, Chen T, Zhou L, et al. Purification and X-ray crystallographic analysis of 7-keto-8-aminopelargonic acid (KAPA) synthase from *Mycobacterium smegmatis*. Acta Crystallogr Sect F: Struct Biol Cryst Commun 2014;F70:1372–5.

Hinds J, Mahenthiralingam E, Kempsell KE, Duncan K, Stokes RW, Parish T, et al. Enhanced gene replacement in mycobacteria. Microbiology 1999;145:519–27.

Keer J, Smeulders MJ, Gray KM, Williams HD. Mutants of *Mycobacterium smegmatis* paired in stationary-phase survival. Microbiology 2000;146:2209–17.

Kitahara T, Hotta K, Yoshida M, Okami Y. Biological studies of amiclenomycin. J Antibiot (Tokyo) 1975;28:215–21.

Lin S, Hanson RE, Cronan JE. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat Chem Biol 2010;9:682–8.

Mann S, Ploux O. 7,8-Diaminoperlargonic acid aminotransferase from *Mycobacterium tuberculosis*, a potential therapeutic target. Characterization and inhibition studies. FEBS J 2006;273:4778–89.

Minnikin DE, Kremer L, Dover LG, Besra GS. The methyl-branched fortifications of *Mycobacterium tuberculosis*. Chem Biol 2002;9:545–53.

Mock DM, Malik MI. Distribution of biotin in human plasma: most of the biotin is not bound to protein. Am J Clin Nutr 1992;56:427–32.

Moss J, Lane MD. The biotin-dependent enzymes. Adv Enzymol Relat Areas Mol Biol 1971;35:321–442.

Okami Y, Kitahara T, Hamada M, Naganawa H, Kondo S. Studies on a new amino acid antibiotic, amiclenomycin. J Antibiot (Tokyo) 1974;27:656–64.

Parish T, Stoker NG. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* tlyA plcABC mutant by gene replacement. Microbiology 2000;146:1969–75.

Ploux O, Marquet A. Mechanistic studies on the 8-amino-7-oxopelargonate synthase, a pyridoxal-5′-phosphate-dependent enzyme involved in biotin biosynthesis. Eur J Biochem 1996;1:301–8.

Ploux O, Breyne O, Carillon S, Marquet A. Slow-binding and competitive inhibition of 8-amino-7-oxopelargonate synthase, a pyridoxal-5'-phosphate-dependent enzyme involved in biotin biosynthesis, by substrate and intermediate analogs. Kinetic and binding studies. Eur J Biochem 1999;259:63–70.

Rengarajan J, Bloom BR, Rubin EJ. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A 2005;102:8327–32.

Salaemae W, Azhar A, Booker GW, Polyak SW. Biotin biosynthesis in *Mycobacterium tuberculosis*: physiology, biochemistry and molecular intervention. Protein Cell 2011;2:691–5.

Sassetti CM, Rubin EJ. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A 2003;100:12989–94.

Shenoy AR, Visweswariah SS. Site-directed mutagenesis using a single mutagenic oligonucleotide and DpnI digestion of template DNA. Anal Biochem 2003;319:335–6.

Said HM. Cell and molecular aspects of human intestinal biotin absorption. J Nutr 2009;139:158–62.

Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol 1990;4:1911–9.

Towbin H, Staehelin T, Gordon J. Electro-phoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A 1979;76:4350–4.

Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, et al. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. Science 2004;305:1466–70.

Walensky LD, Bird GH. Hydrocarbon-stapled peptides: principles, practice, and progress. J Med Chem 2014;15:6275–88.

Webster SP, Alexeev D, Campopiano DJ, Watt RM, Alexeeva M, Sawyer L, et al. Mechanism of 8-amino-7-oxononanoate synthase: spectroscopic, kinetic, and crystallographic studies. Biochemistry 2000;3:516–28.

Woong Park S, Klotzsche M, Wilson DJ, Boshoff HI, Eoh H, Manjunatha U, et al. Evaluating the sensitivity of *Mycobacterium tuberculosis* to biotin deprivation using regulated gene expression. PLoS Pathogens 2011;7:e1002264.

Yu J, Niu C, Wang D, Li M, Teo W, Sun G, et al. MMAR_2770, a new enzyme involved in biotin biosynthesis, is essential for the growth of Mycobacterium marinum in macrophages and zebrafish. Microbes Infect 2011;13:33–41.